

REGRESI LOGISTIK BINER UNTUK MENGETAHUI FAKTOR RISIKO YANG MEMPENGARUHI STROKE NON- HEMORAGIK PADA USIA PRODUKTIF

Arinda Tri Astuti¹⁾, Sri Subanti²⁾, Yuliana Susanti³⁾

^{1) 2) 3)} Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Sebelas Maret
arindatrias3@gmail.com, sri_subanti@yahoo.co.id, yulsusan@yahoo.com

Abstrak

Stroke non-hemoragik merupakan kejadian tersumbatnya pembuluh darah, yang mengakibatkan aliran darah ke otak terhenti. Saat ini stroke non-hemoragik tidak hanya menyerang usia lanjut, namun juga menyerang usia produktif. Tingginya angka kejadian stroke non-hemoragik pada usia produktif merupakan akibat dari ketidakpedulian terhadap faktor risiko yang menyebabkan munculnya kejadian stroke non-hemoragik. Faktor risiko yang berpengaruh terhadap kejadian stroke non-hemoragik, diantaranya jenis kelamin, indeks massa tubuh, High Density Lipoprotein (HDL), Low Density Lipoprotein (LDL), trigliserida, kolesterol total, tekanan darah, dan kadar gula darah. Dua tipe stroke non-hemoragik yang paling dominan adalah stroke trombotik dan stroke embolik. Regresi logistik biner digunakan untuk menganalisis tipe stroke non-hemoragik yang mempunyai skala dikotomis terhadap faktor risiko yang berpengaruh. Penelitian ini bertujuan untuk mengetahui model regresi logistik biner dan besarnya pengaruh faktor risiko yang diduga mempengaruhi kejadian stroke non-hemoragik menggunakan regresi logistik biner. Jumlah sampel penelitian sebanyak 41 pasien. Hasil analisis menunjukkan bahwa kejadian stroke embolik dibandingkan stroke trombotik akan menurun sebesar 1,5404 jika terdapat perubahan dari kadar HDL (≤ 40 mg/dl) ke kadar HDL (> 40 mg/dl) dan pasien yang mempunyai kadar HDL (≤ 40 mg/dl) untuk didiagnosa stroke embolik 4,6667 kali lipat dibanding seseorang yang mempunyai kadar HDL (> 40 mg/dl).

Kata Kunci: Faktor risiko; Regresi logistik biner; Stroke non-hemoragik; Usia produktif

1. PENDAHULUAN

Stroke adalah kondisi yang terjadi ketika pasokan darah ke otak terputus akibat penyumbatan atau pecahnya pembuluh darah, sehingga terjadi kematian sel-sel pada sebagian area di otak (Kemenkes RI, 2016). Hasil Riset Kesehatan Dasar 2018 menyatakan bahwa stroke merupakan salah satu penyakit tidak menular yang mengalami peningkatan prevalensi dari 7% pada tahun 2013 menjadi 10,9% pada tahun 2018 (Kemenkes RI, 2018). Kabupaten Wonogiri merupakan kabupaten dengan prevalensi stroke ketiga tertinggi setelah Kabupaten Boyolali dan Kota Salatiga (Kemenkes RI, 2016). Menurut Dinas Kesehatan Kabupaten Wonogiri (2015), prevalensi stroke meningkat setiap tahunnya dan stroke merupakan penyebab kematian kedua setelah hipertensi. Menurut Mardjono (1981), stroke diklasifikasikan menjadi dua, yaitu stroke hemoragik dan stroke non-hemoragik. Stroke hemoragik adalah stroke yang disebabkan oleh pecahnya pembuluh darah otak, sedangkan stroke non-hemoragik adalah tersumbatnya pembuluh darah yang menyebabkan aliran darah ke otak sebagian

atau keseluruhan terhenti. Menurut Mardjono (1981), dua tipe stroke non-hemoragik yang paling dominan adalah stroke trombotik (57%) dan stroke embolik (15%). Stroke trombotik terjadi karena adanya penggumpalan pembuluh darah ke otak, sedangkan stroke embolik adalah tertutupnya pembuluh arteri oleh bekuan darah.

Penderita stroke meningkat pada usia produktif disebabkan oleh gaya hidup, terutama kebiasaan makan tinggi kolesterol (Dourman, 2013). Kejadian stroke merupakan kejadian yang sangat berbahaya karena dapat mengakibatkan kecacatan bahkan kematian, untuk mengurangi kejadian stroke yang terus meningkat diperlukan upaya pencegahan mengenai apa saja faktor risiko yang memicu kejadiannya. Faktor risiko yang memicu stroke ada dua, yaitu faktor risiko yang dapat dikendalikan dan yang tidak dapat dikendalikan. Faktor risiko yang tidak dapat dikendalikan adalah jenis kelamin. Faktor risiko yang dapat dikendalikan antara lain Indeks Massa Tubuh (IMT), *High-Density Lipoprotein* (HDL), *Low-Density Lipoprotein* (LDL), trigliserida, kolesterol total, tekanan darah, dan Kadar Gula Darah (KGD).

Regresi logistik biner adalah suatu metode analisis data yang digunakan untuk mencari hubungan antara variabel respon (Y) yang bersifat biner dengan variabel prediktor (X) (Hosmer, 2000). Regresi logistik tidak memerlukan asumsi normalitas, homokedastisitas, dan autokorelasi, namun masih memerlukan uji multikolinearitas. Regresi logistik tidak memenuhi asumsi normalitas karena regresi logistik berdistribusi binomial, tidak memenuhi asumsi homokedastisitas karena tergantung oleh peluang rasio (π), dan tidak memenuhi asumsi autokorelasi karena tidak bergantung pada waktu. Regresi logistik biner digunakan untuk menganalisis tipe stroke non-hemoragik berupa data dua kategorik (biner) terhadap faktor risiko yang mempengaruhinya. Oleh karena itu, penelitian ini bertujuan untuk mengetahui model regresi logistik biner dan besarnya pengaruh faktor risiko yang diduga mempengaruhi kejadian stroke non-hemoragik menggunakan regresi logistik biner.

2. METODE PENELITIAN

Penelitian dilakukan di Unit Rawat Inap di bagian Saraf RSUD dr. Soediran Mangun Soemarmo Kabupaten Wonogiri pada bulan Desember 2018. Data yang digunakan merupakan data sekunder. Populasi dari penelitian ini adalah seluruh pasien rawat inap di bagian saraf RSUD dr. Soediran Mangun Soemarmo, sedangkan sampel yang digunakan untuk penelitian ini hanya pasien yang terkena stroke non-hemoragik. Teknik pengambilan sampel pada penelitian ini menggunakan teknik *purposive sampling*. Teknik *purposive sampling* adalah salah satu teknik sampling non-random sampling dimana peneliti menentukan pengambilan sampel dengan cara menetapkan kriteria khusus yang sesuai dengan tujuan penelitian sehingga diharapkan dapat menjawab permasalahan penelitian. Kriteria tersebut adalah pasien dengan diagnosis dokter stroke non-hemoragik pada rentang usia produktif yaitu usia 15-64 tahun dan data rekam medis dan laboratorium yang lengkap. Variabel penelitian yang digunakan dalam penelitian ini terdiri dari variabel respon Y dan variabel prediktor X . Variabel respon Y

adalah tipe stroke non-hemoragik yang terdiri dari 2 kategori, yaitu $Y = 0$ Stroke Trombotik dan $Y = 1$ Stroke Embolik, sedangkan variabel prediktor disajikan pada Tabel 1.

Tabel 1 Variabel Prediktor

Variabel	Kode 1	Kode 2
X_1 : Jenis Kelamin	Laki-Laki	Perempuan
X_2 : IMT	$\leq 25 \text{ kg/m}^2$	$> 25 \text{ kg/m}^2$
X_3 : HDL	$\leq 40 \text{ mg/dl}$	$> 40 \text{ mg/dl}$
X_4 : LDL	$< 130 \text{ mg/dl}$	$\geq 130 \text{ mg/dl}$
X_5 : Trigliserida	$< 200 \text{ mg/dl}$	$\geq 200 \text{ mg/dl}$
X_6 : Kolesterol Total	$< 240 \text{ mg/dl}$	$\geq 240 \text{ mg/dl}$
X_7 : Tekanan Darah	$< 160/95 \text{ mmHg}$	$\geq 160/95 \text{ mmHg}$
X_8 : KGD	$< 200 \text{ mg/dl}$	$\geq 200 \text{ mg/dl}$

Langkah-langkah yang dilakukan untuk melaksanakan penelitian ini sebagai berikut :

1. Mendeskripsikan karakteristik dari variabel tipe stroke non-hemoragik dan variabel prediktor (X_1, \dots, X_8).
2. Melakukan analisis data menggunakan regresi logistik biner, yaitu
 - a. Uji Multikolinearitas

Uji ini digunakan untuk melihat ada atau tidaknya korelasi yang tinggi antara variabel-variabel prediktor dalam suatu model regresi berganda. H_0 menyatakan bahwa nilai tidak terdapat multikolinearitas ($VIF < 10$). H_0 ditolak jika $VIF \geq 10$. Statistik uji VIF sebagai berikut,

$$VIF = \frac{1}{1 - R^2} \tag{1}$$

- b. Estimasi Parameter

Metode yang digunakan untuk mengestimasi parameter menggunakan *Maximum Likelihood Estimator (MLE)*. Prinsip dari metode *MLE* adalah memaksimalkan nilai estimasi β . Berikut merupakan persamaan fungsi *likelihood*,

$$l(\beta) = \prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}}] \tag{2}$$

untuk mengestimasi fungsi *likelihood*, maka persamaan (4) dimaksimalkan, menjadi fungsi *log-likelihood*,

$$L(\beta) = \sum_{i=1}^n y_{1i}g_1(x_i) + y_{2i}g_2(x_i) - \ln(1 + \exp g_1(x_i) + \exp g_2(x_i)) \tag{3}$$

untuk mendapatkan nilai β yang memaksimalkan $L(\beta)$ maka dilakukan turunan terhadap (β) , dengan syarat $\frac{\partial L(\beta)}{\partial \beta_{jk}} = 0$ dan $\frac{\partial^2 L(\beta)}{\partial \beta_{jk} \partial \beta_{jk'}} < 0$ nilai β dapat ditentukan, namun sangat sulit menghitung nilai β secara manual. Oleh karena itu, digunakan metode iterasi dengan komputer untuk mencari solusi nilai β . (Hosmer dan Lameshow, 1989).

c. Uji Rasio *Likelihood*

Pengujian ini dilakukan untuk mengetahui apakah taksiran parameter yang diperoleh berpengaruh signifikan terhadap model atau tidak. Uji signifikansi secara serentak dilakukan dengan uji rasio *likelihood*. Menurut Hosmer dan Lameshow (2000), suatu statistik uji rasio *likelihood* G adalah fungsi L_0 dan L_1 yang berdistribusi *Chi-Square* dengan derajat prediktor p (banyaknya variabel prediktor yang ada dalam model) yang dirumuskan sebagai berikut,

$$G = -2(L_0 - L_1) \tag{4}$$

dengan L_0 adalah *log-likelihood* dari model tanpa prediktor, sedangkan L_1 adalah *log-likelihood* dari model dengan p prediktor. Nilai *log likelihood* dihitung berdasarkan persamaan (3). H_0 menyatakan bahwa $\beta_0 = \beta_1 = \dots = \beta_p = 0$ (Semua variabel prediktor tidak berpengaruh signifikan terhadap variabel respon). H_0 ditolak jika $G > \chi^2_{p,\alpha}$ atau nilai signifikansi $< \alpha$.

d. Uji *Wald*

Uji ini merupakan pengujian parameter secara parsial setiap variabel prediktor terhadap variabel respon. Statistik uji dirumuskan sebagai berikut,

$$W_k = \left(\frac{\widehat{\beta}_k}{SE(\widehat{\beta}_k)} \right)^2 \tag{5}$$

dengan $\widehat{\beta}_k$ menyatakan penduga bagi β_k dan $SE(\widehat{\beta}_k)$ menyatakan penduga galat baku (standar error) bagi β_k . Statistik uji W_k mendekati distribusi *chi-square* dengan derajat prediktor 1. H_0 menyatakan bahwa $\beta_k = 0$ (Variabel prediktor ke $-k$ tidak berpengaruh signifikan terhadap variabel respon) dengan $k = 1, 2, \dots, p$. H_0 ditolak jika $W_k > \chi^2_{1,\alpha}$ atau nilai signifikansi $< \alpha$.

3. Menentukan model regresi logistik biner dari variabel prediktor yang berpengaruh signifikan terhadap stroke non hemoragik.

Model regresi logistik biner dengan p variabel prediktor adalah,

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p)}{1 + \exp(\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p)} \tag{6}$$

Persamaan (2) dapat ditransformasi logit menjadi fungsi linear dari parameternya,

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_px_p \tag{7}$$

(Agresti, 1990)

Setelah mendapatkan model regresi logistik biner, melakukan uji kecocokan model untuk mengevaluasi cocok tidaknya model dengan data. Hipotesis nol (H_0) menyatakan bahwa model cocok dengan data dengan tingkat signifikansi α . Daerah kritis H_0 ditolak jika $D > \chi^2_{(n-p);\alpha}$, dengan statistik uji,

$$D = -2 \sum_{i=1}^k \left\{ y_i \log \left(\frac{n_i \hat{\pi}_i}{y_i} \right) + (n_i - y_i) \log \left(\frac{n_i - n_i \hat{\pi}_i}{n_i - y_i} \right) \right\} \tag{8}$$

(Nugraha, 2013).

4. Menentukan faktor risiko berdasarkan odds rasio (OR) dari setiap variabel prediktor yang mempengaruhi stroke non hemoragik pada usia produktif (15-64 tahun).

Secara umum, OR merupakan sekumpulan peluang yang dibagi oleh peluang lainnya. Nilai OR di definisikan sebagai berikut,

$$OR = \frac{\frac{\pi_1}{(1 - \pi_1)}}{\frac{\pi_0}{(1 - \pi_0)}} = \frac{\exp(\beta_0 + \beta_1)}{\exp(\beta_0)} = \exp \beta_1 \quad (9)$$

(Hosmer dan Lameshow, 2000).

3. HASIL PENELITIAN DAN PEMBAHASAN

3.1 Karakteristik Responden

Analisis deskriptif bertujuan untuk menggambarkan secara umum karakteristik penderita stroke non-hemoragik pada usia produktif di RSUD dr. Soediran Mangoen Sumarso Wonogiri yang dijadikan objek penelitian pada penelitian ini. Objek penelitian yang digunakan adalah penderita stroke non-hemoragik pada usia produktif sebanyak 41 pasien.

Tabel 2 Karakteristik Responden

Variabel		N	%
Y: Tipe Stroke	Stroke Trombotik	27	65.9%
	Stroke Embolik	14	34.1%
X ₁ : Jenis Kelamin	L	22	53,7%
	P	19	46.3%
X ₂ : IMT	≤ 25 kg/m ²	29	70.7%
	> 25 kg/m ²	12	29.3%
X ₃ : HDL	≤ 40 mg/dl	14	34.1%
	> 40 mg/dl	27	65.9%
X ₄ : LDL	< 130 mg/dl	25	61.0%
	≥ 130 mg/dl	16	39.0%
X ₅ : Trigliserida	< 200 mg/dl	36	87.8%
	≥ 200 mg/dl	5	12.2%
X ₆ : Kolesterol total	< 240 mg/dl	36	87.8%
	≥ 240 mg/dl	5	12.2%
X ₇ : Tekanan darah	< 160/95 mmHg	21	51.2%
	≥ 160/95 mmHg	20	48.8%
X ₈ : KGD	< 200 mg/dl	37	90.2%
	≥ 200 mg/dl	4	9.8%

3.2 Uji Multikolinearitas

Uji ini dilakukan sebelum melakukan analisis pemodelan. Nilai VIF didapat dari persamaan (1). Berdasarkan Tabel 3 tidak terdapat variabel prediktor

yang mempunyai nilai VIF >10. Maka dapat disimpulkan bahwa tidak terdapat multikolinieritas dari setiap variabel prediktor.

Tabel 3 Nilai VIF

Variabel	VIF	Variabel	VIF
Jenis Kelamin	1,375	Trigliserida	1,140
IMT	1,198	Kolesterol Total	1,358
HDL	1,110	Tekanan Darah	1,469
LDL	1,215	KGD	1,518

3.3 Uji Rasio Likelihood

Hipotesis nol (H_0) menyatakan bahwa $\beta_0 = \beta_1 = \dots = \beta_8 = 0$ (Semua variabel prediktor tidak berpengaruh signifikan terhadap tipe stroke non-hemoragik). H_0 ditolak jika $G > \chi^2_{8,\alpha} = 15,507$ dengan tingkat signifikansi = 0.05. Nilai G diperoleh dari persamaan (4) sebesar 45,225. Maka dapat disimpulkan bahwa paling tidak terdapat satu variabel prediktor yang berpengaruh terhadap tipe stroke non-hemoragik.

3.4 Uji Wald

Hipotesis nol (H_0) menyatakan bahwa $\beta_k = 0$ (Tidak ada pengaruh antara variabel prediktor ke- k terhadap tipe stroke non-hemoragik), dengan tingkat signifikansi $\alpha = 0.05$. Daerah kritis H_0 ditolak jika $W_k > \chi^2_{(1,\alpha)}$ atau H_0 ditolak jika $Sig. < 0,05$.

Tabel 3 Nilai Uji Wald

Variabel	β	SE	Wald	df	Sig.
Intersep	0,3914	0,9393	0,1739	1	0,6769
Jenis Kelamin Perempuan	0,4660	0,8240	0,3204	1	0,5717
IMT > 25 kg/m ²	-0,4098	0,8917	0,2116	1	0,6458
HDL > 40 mg/dl	-1,7759	0,8065	4,8488	1	0,0277**
LDL >= 130 mg/dl	0,6228	0,7797	0,6384	1	0,4244
Trigliserida >= 200 mg/dl	-0,8410	1,3721	0,3758	1	0,5399
Kolesterol Total >= 240 mg/dl	1,3164	1,2568	1,0962	1	0,2949
Tekanan darah >= 160/95 mg/dl	-0,7627	0,8424	0,8190	1	0,3652
KGD >= 200 mg/dl	-0,3801	1,3604	0,0778	1	0,7799

** signifikan 5%

Nilai *Wald* pada Tabel 4 diperoleh dari persamaan (5). Kesimpulannya adalah terdapat pengaruh yang signifikan antara variabel HDL terhadap tipe stroke non-hemoragik. Oleh karena itu dilakukan analisis kembali yang hanya melibatkan variabel HDL ke dalam persamaan. Namun, uji multikolinieritas dan uji rasio *likelihood* tidak dilakukan karena hanya ada satu variabel saja yang diuji. Kemudian dilakukan uji *Wald* untuk mengetahui nilai koefisien β dan nilai OR pada persamaan (9).

Tabel 4 Nilai Odds Rasio

Variabel	β	Std.Error	Wald	df	$\exp(\beta)$	Sig.
Intersep	0,2877	0,5401	0,2841	1	0,2143	0,5943
HDL > 40 mg/dl	-1,5404	0,7113	4,6916	1	0,46667	0,0303**

** signifikan 5%

Selanjutnya, dibentuk fungsi linear dari parameter β seperti persamaan (7) diperoleh dari Tabel 5,

$$g(x) = \ln \left[\frac{P(Y = Stroke Embolik)}{P(Y = Stroke Trombotik)} \right] = \beta_0 + \beta_3 x_3$$

$$g(x) = \ln \left[\frac{P(Y = Stroke Embolik)}{P(Y = Stroke Trombotik)} \right]$$

$$= 0,2877 - 1,5404 x_3$$

$$g(x) = \ln \left[\frac{P(Y = Stroke Embolik)}{P(Y = Stroke Trombotik)} \right]$$

$$g(x) = 0,2877 - 1,5404 HDL (> 40 mg/dl) \tag{10}$$

Interpretasi dari persamaan (10) yaitu kejadian stroke embolik dibandingkan stroke trombotik akan menurun sebesar 1,5404 jika terdapat perubahan dari kadar HDL (kolesterol baik) rendah ($\leq 40 mg/dl$) ke kadar HDL (kolesterol baik) tinggi ($> 40 mg/dl$). Model regresi logistik biner sebagai berikut,

$$\pi(x) = \frac{e^{0,2877-1,5404 x_3}}{1 + e^{0,2877-1,5404 x_3}} \tag{11}$$

3.5 Uji Kecocokan Model

Hipotesis nol uji Devians menyatakan bahwa model regresi logistik biner layak digunakan dengan tingkat signifikansi $\alpha = 0,05$. Daerah kritis H_0 ditolak jika $D > \chi^2_{(40;0,05)} = 55,7585$.

Tabel 5 Nilai Devians

Devians	Sig.
47,7256	0,1594

Nilai Devians pada Tabel 6 didapat dari persamaan (8), yaitu $D = 47,7256 < \chi^2_{(40;0,05)} = 55,7585$. Maka, dapat disimpulkan bahwa H_0 diterima. Sehingga model regresi logistik biner layak digunakan.

3.6 Odds Rasio (OR)

Odds rasio (OR) digunakan untuk menjelaskan peluang dari masing-masing variabel penjelas terhadap variabel respon. Nilai OR didapat dari persamaan (9).

Tabel 6 Nilai Odds Rasio Variabel HDL

Variabel Prediktor	Tipe Stroke Non-hemoragik	
	Stroke Embolik - Stroke Trombotik	
	Exp(B)	
HDL > 40 mg/dl	0,2143	
HDL ≤ 40 mg/dl	4,6667	

Pada variabel HDL, hasil menunjukkan bahwa untuk prediksi diagnosa stroke embolik dibandingkan dengan stroke trombotik, kecenderungan seseorang yang mempunyai kadar HDL tinggi (> 40 mg/dl) untuk didiagnosa stroke embolik 0,2143 kali lipat dibanding seseorang yang mempunyai kadar HDL rendah (≤ 40 mg/dl). Interpretasi lainnya adalah kecenderungan seseorang yang mempunyai kadar HDL rendah (≤ 40 mg/dl) untuk didiagnosa stroke embolik 4,6667 kali lipat dibanding seseorang yang mempunyai kadar HDL tinggi (> 40 mg/dl).

4. SIMPULAN

Hasil analisis dan pembahasan dapat disimpulkan bahwa kejadian stroke embolik dibandingkan stroke trombotik akan menurun sebesar 1,5404 jika terdapat perubahan dari kadar HDL (≤ 40 mg/dl) ke kadar HDL (> 40 mg/dl) dan pasien yang mempunyai kadar HDL (≤ 40 mg/dl) untuk didiagnosa stroke embolik 4,6667 kali lipat dibanding seseorang yang mempunyai kadar HDL (> 40 mg/dl).

5. DAFTAR PUSTAKA

- Agresti, A. (1990). *Categorical Data Analysis*. USA: John Wiley & Sons Inc.
- Dinas Kesehatan Kabupaten Wonogiri. (2015). *Profil Kesehatan Kabupaten Wonogiri Tahun 2015*. Wonogiri: Dinkes Kabupaten Wonogiri.
- Dourman, K. (2013). *Waspada Stroke Usia Muda*. Jakarta: Cerdas Sehat.
- Hosmer, D.W, and Lemeshow, S. (2000). *Applied Logistic Regression*. USA: John Willey, New York.
- Kementrian Kesehatan Republik Indonesia. (2013). *Badan Penelitian dan Pengembangan Kesehatan*. Jakarta: Riset Kesehatan Dasar 2013.
- _____. (2016). *Badan Penelitian dan Pengembangan Kesehatan*. Jakarta: Riset Kesehatan Dasar 2016.
- _____. (2018). *Potret Sehat Indonesia dari Riskesdas 2018*. Diakses dari <http://www.depkes.go.id/article/view/18110200003/potret-sehat-indonesia-dari-riskesdas-2018.html>
- Mardjono, M. (1981). *Neurologi Klinis Dasar*. Jakarta: Dian Rakyat.
- Nugraha, J. (2013). *Pengantar Analisis Data Kategorik*. Jogjakarta: Deepublish.